PROV-O Provenance Traces From Agent-based Social Simulation

Edoardo Pignotti
Computing Science
University of Aberdeen
Aberdeen AB24 3UE
e.pignotti@abdn.ac.uk

Gary Polhill
The James Hutton Institute
Craigiebuckler
Aberdeen AB15 8QH
gary.polhill@hutton.ac.uk

Peter Edwards
Computing Science
University of Aberdeen
Aberdeen AB24 3UE
p.edwards@abdn.ac.uk

1. SUMMARY

In this paper we describe a number of PROV-O provenance traces generated by the OBIAMA simulation environment [1]. OBIAMA stands for Ontology-Based Integrated Action Modelling Arena. It is a toolkit designed to facilitate discrete-event modelling using OWL ontologies as the medium in which the structure and state of the model are represented, as opposed to the object-oriented programming framework (classes and objects), as is more commonly the case. Whilst ultimately, since all computer modelling ends up as a series of machine language instructions on a CPU, it matters very little how things are represented, there are certain reasons why OWL ontologies have benefits over OO as a basis for representation. The case for transparency is made by [2]. However, there are other benefits, not least in assisting with integrating modular components in a way that allows sanity checking of ontological coherence.

OBIAMA has a facility to record provenance about which actions performed by agents in the simulation model caused the assertion or retraction of axioms in the ontological model describing simulation state. OBIAMA can record provenance using different representation models: OPM[1], OPMV[2] and more recently PROV-O[3].

In OBIAMA one of the key questions is where axioms have come from. Provenance should thus be able to answer the question of which actions were responsible for each axiom. The diagram in Figure 1 shows how PROV-O is used to represent provenance within an OBIAMA simulation.

We have uploaded to ProvBench a number of provenance traces generated by a sample OBIAMA model varying the number of agents and timesteps e.g. in the WERCM-provo-40-20-provo.owl file, 40 represents the number of agents in the simulation model and 20 the number of timesteps.

[1]http://openprovenance.org/
[2]http://open-biomed.sourceforge.net/opmv/ns.html
[3]http://www.w3.org/TR/prov-o/

Data Format	RDF
Data Model	PRO-O
Size	Between 2 and 81 Mb
Tool used for generating provenance	OBIAMA
Domain	Social Simulation
Submission group	OBIAMA
License	cc-by-sa

Table 1: Information About the PROV-O Traces.

2. COVERAGE OF PROV

We now describe in more details how provenance is characterised in a OBIAMA model using the PROV-O model.

Agent : Agents in the model are software entities that are responsible for a change in the model's state. In OBIAMA, all actions must be associated with an agent.

Activity : There are two subclasses of activity. One subclass is an Action e.g. model initialisation, the other is assertion or retraction which describe the process of asserting or retracting OWL axioms in the model's state.

Entity : OWL Axioms in the model state ontology.

wasGeneratedBy : Assertions generate OWL Axioms.

used : Retractions use the retracted Axiom.

wasAttributedTo : Axioms are attributed to the Agent that caused the change.

wasAssociatedWith : Agents are associated with the Actions that they perform.

wasInformedBy : Inferences between assertions and retractions based on the Action performed by an Agent.

wasDerivedFrom : Axioms were derived from other axioms if the Actions used them.

3. EXAMPLE PROVENANCE QUERIES

One of the important aspects of the provenance in OBIAMA is being able to reconstruct the state of a simulation at any given timestep without having to re-run the simulation.

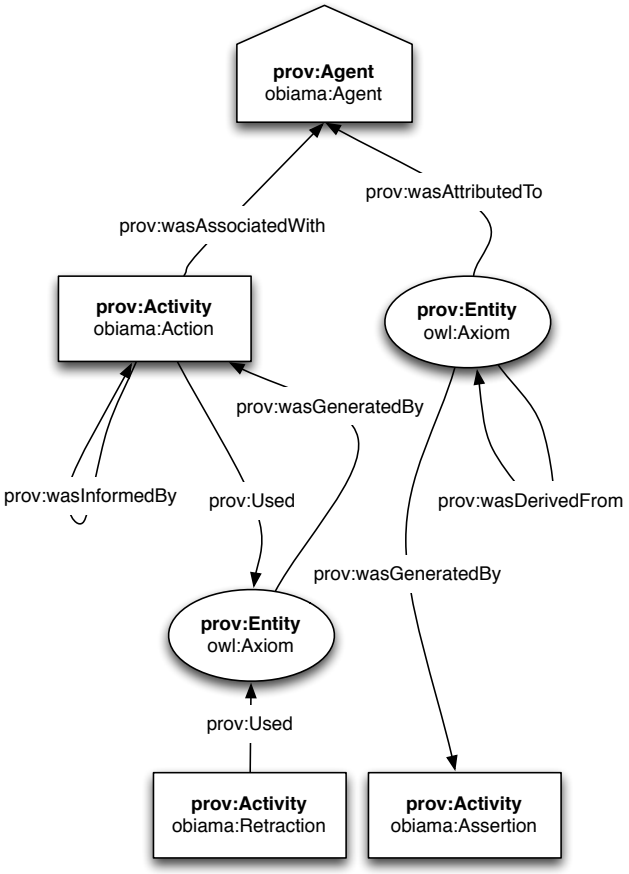

Figure 1: How provenance is defined in OBIAMA.

[Q1] *What axioms are associated with the final state of a simulation model?*

[Q2] *What are the intermediate states?*

Provenance in OBIAMA can be used to identify what activities, agents and entities have been involved in the production of an axiom in the model. It is particularly useful to determine if there are distinct 'spheres' of activity in the model in which, for example, a group of asserted axioms is disconnected form the rest of the provenance record.

[Q3] *How did the model come to produce axiom X?*

Debugging and system testing could also benefit by the use of provenance as it will be possible to understand how the model came to produce a particular axiom. System testing is the process of verifying that the model behaves as expected. When it does not do so, we need to know why, and that is where provenance could be useful too. For debugging purposes it will also be useful to know why a particular axiom does NOT appear in a particular state (e.g. which actions retract it).

[Q4] *Why does axiom X do not in a state S?*

[Q5] *Which actions in the simulation retracted axiom X?*

Investigating emergent properties in the model could be facilitated by exploring the provenance across multiple simulation runs. It is important in this kind of exploration to determine if in some runs a phenomenon occurs but in other runs it does not.

[Q6] Given a set of axioms describing a phenomenon in the simulation, identify the provenance paths leading to those axioms?

4. CONCLUSIONS AND FUTURE WORK

We expect to continue to produce provenance traces from a number of different simulation models and experiments. It should be noted that to date we have submitted only submitted the provenance traces of individual model runs, a typical simulation experiment (a parameter exploration study) might involve hundreds of simulation runs and might be required to query provenance across different runs.

One issue for modelling environments such as OBIAMA is that the reasoner may be used to infer axioms about the model's state using axioms generated by activities. This means the provenance in OBIAMA ideally needs to record when an axiom has been added by the reasoner, and, ideally, which axioms it used to generate them. Future work will include capturing the provenance of the reasoning process in OBIAMA.

5. REFERENCES

[1] G. Polhill, N. Gotts, N. Sánchez-Maroño, E. Pignotti, Ó. Fontenla-Romero, M. Rodríguez-García, A. Alonso-Betanzos, P. Edwards, and T. Craig. An ontology-based design for modelling case studies of everyday pro-environmental behaviour in the workplace. *International Congress on Environmental Modelling and Software, Leipzig, Germany, 1-5 July 2012.*, 2012.

[2] J.G. Polhill and N.M. Gotts. Ontologies for transparent integrated human-natural system modelling. *Landscape ecology*, 24(9):1255–1267, 2009.